

Michał ROGALEWICZ*, Agnieszka KUJAWIŃSKA *

WSPOMAGANIE DECYZJI ZAKUPOWYCH W BRANŻY SPAWALNICZEJ ZA POMOCĄ METODY k -ŚREDNICH

DOI: 10.21008/j.0239-9415.2016.070.15

Podejmowanie decyzji w przedsiębiorstwie wiąże się często z wyborem najlepszego rozwiązania na podstawie wielu kryteriów opisujących analizowany problem. Z tego punktu widzenia można go nazwać wielokryterialnym problemem decyzyjnym. W artykule przedstawiono zastosowanie jednej z metod wspomaganie decyzji – analizy skupień metodą k -średnich – w doborze materiałów dodatkowych do procesu spawania metodą SAW. Dokonano podziału na skupienia, uwzględniając dwa kryteria doboru ich początkowych centrów, porównano oba warianty, a na końcu scharakteryzowano szczegółowo grupy wyodrębnione za pomocą jednego z nich. Wybrane podejście do analizy skupień okazało się przydatne we wspomaganie decyzji dotyczących zakupów w branży spawalniczej.

Słowa kluczowe: wspomaganie decyzji, analiza skupień, metoda k -średnich

1. WPROWADZENIE

Skuteczne zarządzanie systemem produkcyjnym w przedsiębiorstwie wiąże się z potrzebą podejmowania decyzji związanych z ryzykiem oraz niepewnością, ze względu na dużą liczbę możliwości i towarzyszących im kryteriów. Na każdym ze szczebli struktury organizacyjnej pracownicy rozwiązują problemy, często polegając jedynie na własnej wiedzy, doświadczeniu i intuicji. Ponieważ skutki błędnych lub nieoptymalnych decyzji wiążą się z poniesieniem kosztów, coraz częściej w przedsiębiorstwach dostrzega się potrzebę wspomaganie decyzji i zastosowania do tego naukowych metod. Jednym z wielokryterialnych problemów decyzyjnych, do którego rozwiązania można użyć takich metod, jest zakup zasobów niezbędnych w procesach wytwarzania. Duża liczba alternatyw oraz informacji na temat każdej

* Politechnika Poznańska, Wydział Budowy Maszyn i Zarządzania.

z nich powoduje dezorientację decydenta i może zniechęcić go do zakupu lub spowodować, że jego wybór nie będzie optymalny.

Przykładem takiego problemu jest podejmowanie decyzji o zakupie tzw. materiałów dodatkowych (topników) do spawania metodą SAW (ang. *submerged arc welding* – spawanie łukiem krytym).

Topniki wspomagają stabilne jarzenie łuku elektrycznego, chronią roztopiony metal przed wpływem otaczającej atmosfery, umożliwiają rafinowanie stopiwa, wprowadzają do niego składniki stopowe oraz kształtują powstającą spoinę, a więc pełnią podobną funkcję jak otulina elektrody przy ręcznym spawaniu łukowym. Technologia SAW pozwala na zastosowanie wysokiego natężenia prądu oraz dużej prędkości spawania. Rodzaj topnika ma znaczący wpływ na właściwości użytkowe, przede wszystkim na udarność spoiny. Wyróżnia się cztery rodzaje topników ze względu na skład chemiczny: kwaśne, obojętne, zasadowe i wysokozasadowe. Topniki kwaśne i obojętne wykazują niższą temperaturę topnienia niż stopiwo, czego efektem są bardzo dobre właściwości spawalnicze, ale wiąże się z tym również występowanie większej ilości wtrąceń obniżających udarność spoiny. Zastosowanie topników zasadowych i wysokozasadowych skutkuje z kolei uzyskaniem spoiny o wysokiej czystości, co w znaczący sposób zwiększa udarność w niskiej temperaturze (Houldcroft, 2014).

Wybranie konkretnego rodzaju topnika spośród szerokiej ich gamy dostępnej obecnie na rynku nie jest trywialne. Duża liczba kryteriów, jakimi musi się kierować inżynier dobierający materiały dodatkowe do procesu spawania (kryteria te to m. in. wskaźnik zasadowości Boniszewskiego, zawartość wybranych pierwiastków stopowych, właściwości mechaniczne spoiny i cena) sprawia, że jest to jeden z kluczowych problemów w jego pracy. Jak już wspomniano, im większa wiedza i doświadczenie decydenta, tym większe jest prawdopodobieństwo podjęcia prawidłowej decyzji, jednak nie zawsze decyzja ta będzie optymalna.

W takiej sytuacji pomocne mogą się okazać metody wielowymiarowej analizy i odkrywania wiedzy z danych, określane wspólnym mianem – *Data Mining* (Hastie, Tibshirani, Friedman, 2001; Han, Kamber, 2006). Spośród zadań, wykonywanych z ich zastosowaniem (m. in. klasyfikacja i wyznaczanie funkcji regresji (Cao, 2010)) szczególnie przydatne podczas dokonywania oceny wielu obiektów (w tym przypadku topników) opisanych dużą liczbą kryteriów decyzyjnych może się okazać ich grupowanie (ang. *clustering* (Jain, Murty, Flynn, 1999)). Polega ono na utworzeniu skupisk obiektów, które są do siebie najbardziej „podobne” i zastępowalne w ramach wyodrębnionej grupy. Takie podejście może znacznie ułatwić decydentowi wybór odpowiedniego topnika.

W obszarze *Data Mining* zadanie grupowania jest realizowane przez wyodrębniony zbiór metod, a ich stosowanie nazywa się analizą skupień.

2. ANALIZA SKUPIEŃ

Analiza skupień ma za zadanie przydział zbioru analizowanych obiektów/przypadków do jednorodnych grup obiektów (skupień) na podstawie opisujących je kryteriów. Powstałe skupienia powinny jak najbardziej różnić się od siebie, ale jednocześnie obiekty będące elementami tej samej grupy powinny wykazywać jak największe podobieństwo. Metody należące do obszaru analizy skupień różnią się m.in. sposobem wyodrębniania skupień i stosowanymi miarami podobieństwa między obiektami/i skupieniami (Abbas, 2008).

Wyróżnia się dwa sposoby tworzenia skupień. Pierwszy polega na iteracyjnym tworzeniu grup w taki sposób, że grupa na danym poziomie składa się z grup otrzymanych we wcześniejszych krokach. Jest to tzw. podejście hierarchiczne (ang. *hierarchical clustering*) (Murtagh, 2012). Obejmuje ono dwie techniki: podziałowe i aglomeracyjne. Metody podziałowe zakładają, że na początku analizy wszystkie obiekty w analizowanym zbiorze stanowią jedno skupienie, a następnie ulega ono dzieleniu na mniejsze grupy. Proces jest powtarzany aż do momentu, gdy wszystkie analizowane elementy utworzą swoje własne, jednoelementowe skupienia. Zdecydowanie częściej niż techniki podziałowe stosuje się techniki aglomeracyjne. Na początku ich użycia każdy obiekt stanowi oddzielne skupienie. Następnie dochodzi do łączenia jednoelementowych skupień w nowe skupienia, które zawierają elementy najbardziej do siebie podobne. Proces powtarza się aż do momentu uzyskania jednego skupienia zawierającego wszystkie elementy zbioru. Efektem analizy dokonanej według podejścia hierarchicznego jest powstanie dendrogramu, czyli hierarchii drzewa analizowanych obiektów. Umożliwia ono zaobserwowanie zależności między powstałymi skupieniami.

Do drugiej grupy metod analizy skupień należą metody niehierarchiczne (ang. *non-hierarchical clustering*) (Berkhin, 2006). W odróżnieniu od metod hierarchicznych w ich ramach wyodrębnia się z góry określoną przez decydenta liczbę skupień, ale w taki sposób, że nie tworzą one hierarchii, a są rozłączne (oznacza to, że żadne skupienie nie jest podskupieniem innego). Ważnym parametrem jest tutaj liczba tworzonych grup, którą często wyznacza się metodą prób i błędów. Wytypowanie zbyt dużej liczby skupień powoduje powstanie kilku skupień bardzo podobnych. Wskazanie zbyt małej ich liczby może doprowadzić do dużego zróżnicowania wewnętrznego każdej grupy.

Istnieje kilka elementów decydujących o rzetelności przeprowadzanej analizy skupień. Są to m. in.:

- sposób akwizycji danych oraz ich jakość,
- odpowiednie przygotowanie danych do analizy, wiążące się m.in. z eliminacją wpływu wartości odstających, redundantnych kryteriów, braku danych,
- właściwy dobór parametrów, które mogą być różne w zależności od zastosowanej metody.

Ponieważ w analizie skupień grupowania obiektów oraz skupień dokonuje się na podstawie ich podobieństwa, szczególnie istotne jest ustalenie odpowiedniej metryki odległości pomiędzy nimi. W literaturze przedmiotu (Popat, Emmanuel, 2014) można znaleźć szereg różnych metryk; najczęściej stosowane to odległości euklidesowa, Czebyszewa i miejska. Inne parametry ustalane przez użytkownika to np. sposób wiązania ze sobą obiektów i skupień w przypadku metod aglomeracyjnych hierarchicznych (wyróżnia się tu m.in. metody pojedynczego wiązania, pełnego wiązania lub średnich połączeń, ale efektywna wydaje się oparta na analizie wariancji metoda Warda) lub wspomniana już początkowa liczba skupień w metodach niehierarchicznych. Jak już wspomniano, parametry mogą być dobierane na podstawie doświadczenia decydenta lub metodą prób i błędów, ale można do tego zastosować również zaawansowane techniki *Data Mining*, takie jak V-krotna walidacja krzyżowa (Popat, Emmanuel, 2014).

3. CEL BADAŃ I SPOSÓB PRZYGOTOWANIA DANYCH DO ANALIZY

3.1. Cel badań i akwizycja danych

Celem prowadzonych badań była ocena możliwości i zasadności zastosowania wybranej metody analizy skupień w problemie podejmowania decyzji dotyczącej zakupu topników do spawania metodą SAW w wybranym przedsiębiorstwie branży metalowej. Wyniki analizy mają wspomóc inżyniera procesu, którego dotychczasowe wybory nie zawsze były optymalne ze względu na brane pod uwagę kryteria. Autorzy niniejszego artykułu do analizy wybrali niehierarchiczną metodę *k*-średnich celem utworzenia jednorodnych, rozłącznych skupisk najbardziej „podobnych do siebie” obiektów.

Praca badawcza wiązała się ze zgromadzeniem danych dotyczących topników z ofert handlowych pochodzących od pięciu producentów. Kryteriami wyboru topnika były: wskaźnik zasadowości Boniszewskiego, właściwości spoiny powstałej po procesie spawania (zawartość pierwiastków – węgla, krzemu, manganu, molibdenu, chromu i niklu oraz wytrzymałość, granica plastyczności, wydłużenie) oraz cena.

Warto zauważyć, że na właściwości spoiny po spawaniu ma wpływ nie tylko topnik, ale również zastosowany drut topliwy. Ponieważ istnieje silna zależność między tymi dwoma dodatkowymi materiałami do spawania metodą SAW (użycie tego samego topnika z różnymi drutami skutkuje uzyskaniem spoin o różnych właściwościach mechaniczno-chemicznych), w badaniach przeanalizowano właściwości topnika oraz spoiny z rozróżnieniem drutu topliwego, który został użyty w procesie spawania. Stąd wstępnie uzyskana liczba kombinacji topników z drutami topliwymi wyniosła 274 (fragment zbioru danych – tab. 1).

Tabela 1. Fragment zbioru danych

Topnik	WZB	C [%]	Si [%]	Mn [%]	Mo [%]	Cr [%]	Ni [%]	Rm [MPa]	Re [MPa]	A5 [%]	Cena [PLN/kg]
A1	2,6	0,07	0,15	0,5	0	0	0	445	355	26	8,93
A2	2,6	0,08	0,35	1	0	0	0	520	440	30	8,93
A3	2,6	0,06	0,25	1	0,5	0	0	570	480	26	8,93
...											
B1	1,7	0,04	0,15	0,9	0	0	0	450	360	24	BD
B2	1,7	0,04	0,15	0,9	0,5	0	0	550	430	20	BD
B3	1,7	0,04	0,2	1,3	0	0	0	500	400	24	BD
...											
C1	1,1	0,05	0,25	1	0	0	0	480	360	30	6,88
C2	1,1	0,06	0,3	1,3	0	0	0	490	390	33	6,88
C3	1,1	0,1	0,3	1,2	0	0	0	510	430	32	6,88
...											
E107	0,4	0,04	0,6	1,3	0	0	0,7	550	470	22	10,63
E108	0,4	0,04	0,6	1,1	0	0	0	520	400	22	10,63
E109	0,4	0,04	0,6	1,3	0	0	0	520	400	22	10,63
...											

Oznaczenia: Topnik – kod producenta; WZB – wskaźnik zasadowości według Boniszewskiego; C – węgiel; Si – krzem; Mn – mangan, Cr – chrom, Ni – nikiel, Rm – wytrzymałość, Re – granica plastyczności, A5 – wydłużenie.

3.2. Przygotowanie danych do analizy

Analiza skupień jest bardzo czuła na jakość danych, które są w niej wykorzystane. Dlatego szczególnie ważne jest odpowiednie ich przygotowanie przed rozpoczęciem obliczeń. W pierwszym kroku autorzy wyeliminowali rekordy w bazie danych zawierające obserwacje odstające oraz braki danych. Ponadto na podstawie badania zmienności cech (za pomocą współczynnika zmienności) wytypowano kryteria o największej zdolności dyskryminacyjnej (tab. 2).

Tabela 2. Współczynnik zmienności dla cech diagnostycznych użytych w analizie

Cecha diagnostyczna	WZB	C	Si	Mn	Mo	Cr	Ni	Rm	Re	A5	Cena
Współczynnik zmienności ω [%]	46,3	42,1	55,3	60,5	197,5	152,1	226,8	15,1	18,6	20,9	31,3

Kolejnym etapem przygotowania bazy danych do właściwej analizy skupień było zbadanie współliniowości kryteriów opisujących kombinację topnik–drut topliwy.

Na podstawie analizy korelacji wykazano wysoką i istotną korelację między kilkoma kryteriami oceny topników (tab. 3 – istotne korelacje wyróżniono czcion-

ką pogrubioną). W jej wyniku zdecydowano się na wyeliminowanie tego z dwóch najsilniej skorelowanych kryteriów, które ma mniejsze zdolności dyskryminacyjne – wytrzymałości na rozciąganie.

Tabela 3. Analiza korelacji między kryteriami

	WZB	% C	% Si	% Mn	% Mo	% Cr	% Ni	Rm [N/mm ²]	(Re 0,2%) [N/mm ²]	(A5) [%]	Cena topnika [pln/kg]
WZB	1,00	0,14	-0,56	-0,10	0,08	-0,02	0,05	0,09	0,13	-0,12	0,04
% C	0,14	1,00	-0,05	-0,09	-0,30	-0,30	-0,24	-0,10	0,11	-0,26	-0,19
% Si	-0,56	-0,05	1,00	0,21	-0,15	-0,05	-0,10	0,04	0,08	0,14	0,12
% Mn	-0,10	-0,09	0,21	1,00	0,04	0,18	0,09	0,05	-0,07	0,28	0,07
% Mo	0,08	-0,30	-0,15	0,04	1,00	0,59	0,68	0,51	0,12	0,25	0,17
% Cr	-0,02	-0,30	-0,05	0,18	0,59	1,00	0,65	0,41	-0,10	0,58	0,40
% Ni	0,05	-0,24	-0,10	0,09	0,68	0,65	1,00	0,32	-0,11	0,49	0,26
Rm [N/mm ²]	0,09	-0,10	0,04	0,05	0,51	0,41	0,32	1,00	0,80	0,00	0,20
(Re 0,2%) [N/mm ²]	0,13	0,11	0,08	-0,07	0,12	-0,10	-0,11	0,80	1,00	-0,34	0,02
(A5) [%]	-0,12	-0,26	0,14	0,28	0,25	0,58	0,49	0,00	-0,34	1,00	0,38
Cena topnika [pln/kg]	0,04	-0,19	0,12	0,07	0,17	0,40	0,26	0,20	0,02	0,38	1,00

Ostatnim etapem przygotowania danych do zastosowania analizy skupień było dokonanie standaryzacji wszystkich kryteriów. Miało to na celu wyeliminowanie wpływu na wynik tych zmiennych, które w danych jednostkach osiągają duże wartości (mogłyby one zdominować miarę odległości).

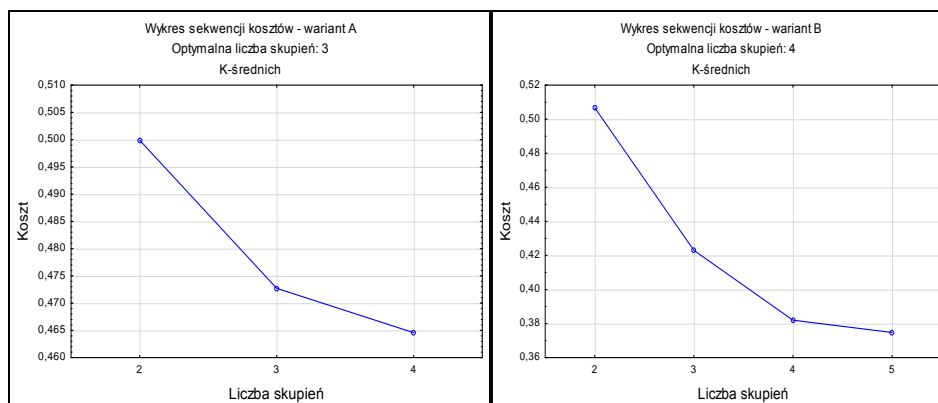
W wyniku akwizycji i przygotowania danych do analizy w finalnym zbiorze danych znalazło się 212 przypadków (kombinacji topników i drutu topliwego oraz właściwości spoiny) oraz 10 cech diagnostycznych.

4. ANALIZA DANYCH

Na tak przygotowanym zbiorze danych dokonano analizy skupień metodą *k*-średnich w programie Statistica. Ze względu na brak apriorycznej wiedzy dotyczącej liczby i centrów skupień wybrano dwa warianty:

- centra maksymalizujące wstępne odległości skupień – wariant A,
- losowy wybór *N* obserwacji jako centra skupień – wariant B.

Dla obu wariantów zastosowano 10-krotną walidację krzyżową, która posłużyła do doboru optymalnej ze względu na ustalone kryterium kosztu liczby skupień. W przypadku A program zaleca trzy skupienia, a w przypadku B – cztery (rys. 1).



Rys. 1. Wykres sekwencji kosztów dla wariantów A i B

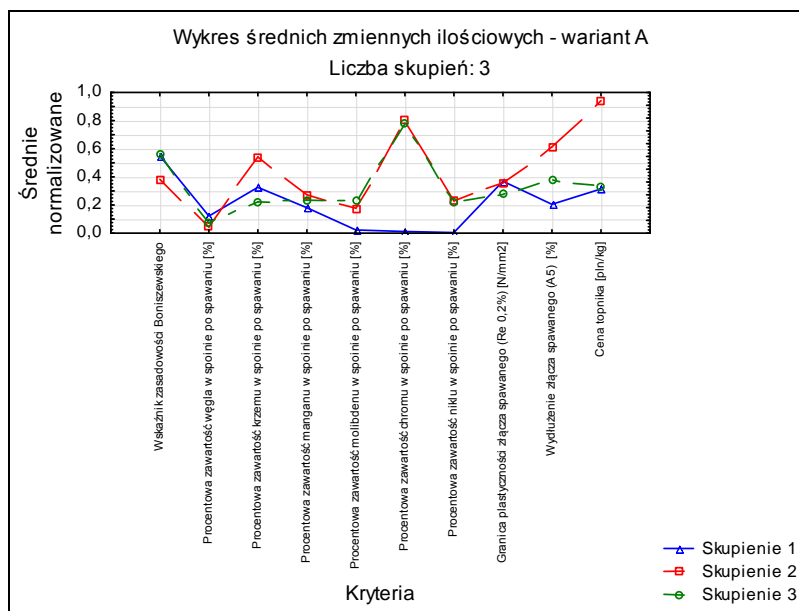
Kolejnym etapem analizy było sprawdzenie istotności kryteriów w aspekcie różnicowania skupień. W tym celu przeprowadzono analizę wariancji, której wyniki przedstawiono w tabeli 4.

Tabela 4. Analiza wariancji dla wariantu A i B (oprac. własne)

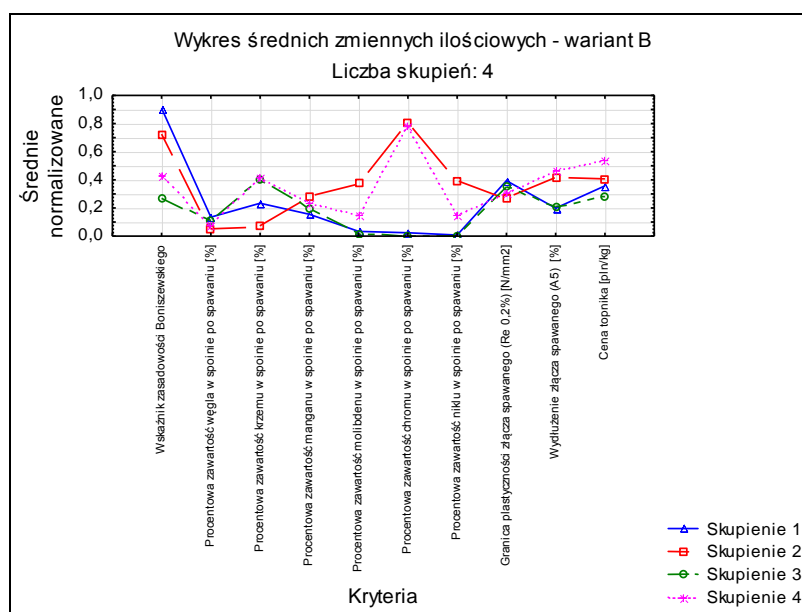
Kryterium	Wariant A		Wariant B	
	wartość statystyki testowej F	poziom p	Wartość statystyki testowej F	poziom p
WZB	2,384	0,0947	251,849	0,0000
% C	13,697	0,0000	10,141	0,0000
% Si	21,670	0,0000	35,329	0,0000
% Mn	6,070	0,0027	5,958	0,0006
% Mo	52,262	0,0000	58,578	0,0000
% Cr	3330,360	0,0000	2237,819	0,0000
% Ni	71,573	0,0000	87,446	0,0000
(Re 0,2%) [N/mm ²]	4,507	0,0121	3,131	0,0266
(A5) [%]	105,101	0,0000	47,583	0,0000
Cena topnika [pln/kg]	178,072	0,0000	16,800	0,0000

Na podstawie analizy wykazano, że w obu przypadkach dziewięć z dziesięciu kryteriów okazało się istotnych (wartości pogrubione w tab. 4 – poziom $p < 0,05$), tzn. wartość średnia co najmniej jednego skupienia istotnie odbiegała od pozostałych. Tylko wskaźnik zasadowości Boniszewskiego wykazywał istotność w przypadku B, ale w przypadku A nie różnicował istotnie skupień. Poniżej przedstawiono wykresy wartości średnich kryteriów dla poszczególnych skupień (w jednost-

kach standaryzowanych) dla wariantów A i B (rys. 2 i 3) oraz tabelę z wartościami średnimi wyrażonymi w oryginalnych jednostkach (tab. 5).



Rys. 2. Wykres średnich dla wariantu A (3 skupienia)



Rys. 3. Wykres średnich dla wariantu B (4 skupienia)

Tabela 5. Analiza wariancji dla wariantów A i B

Kryterium	Wariant A			Wariant B			
	skupienie 1	skupienie 2	skupienie 3	skupienie 1	skupienie 2	skupienie 3	skupienie 4
WZB	1,94	1,47	1,98	2,94	2,43	1,15	1,58
% C	0,07	0,03	0,04	0,07	0,03	0,06	0,04
% Si	0,39	0,65	0,27	0,28	0,09	0,48	0,50
% Mn	1,27	1,91	1,64	1,09	1,93	1,41	1,62
% Mo	0,23	1,77	2,36	0,34	3,77	0,14	1,50
% Cr	0,30	21,06	20,20	0,63	20,79	0,05	20,28
% Ni	0,33	16,95	16,34	0,57	29,00	0,14	10,99
(Re 0,2%) [N/mm ²]	469,77	464,41	426,78	481,21	423,68	460,77	443,02
(A5) [%]	24,56	37,65	30,31	24,38	31,16	24,70	32,84
Cena topnika [pln/kg]	10,16	20,12	10,38	10,85	11,63	9,63	13,68
Liczba przypadków	150	17	45	66	19	84	43

Z porównania obu wariantów wynika, że oba podziały istotnie się różnią. Wariant A składa się z trzech skupień, których centra położone są średnio dalej od siebie (tab. 6), ale skutkuje to również większym zróżnicowaniem wewnętrznym w ramach poszczególnych grup (np. w liczącym 150 kombinacji topnika i drutu topliwego skupieniu 1). Wariant B zawiera cztery skupienia, ale położone bliżej w przestrzeni kryteriów. Są one również mniej zróżnicowane wewnętrznie. Porównanie obiektów należących do skupień wyodrębnionych w ramach wariantów A i B może wzbogacić analizę skupień i dostarczyć dodatkowej wiedzy o badanym zjawisku.

Tabela 6. Odległości pomiędzy centrami skupień w wariantach A i B

	Wariant A			Wariant B				
	S1	S2	S3	S1	S2	S3	S4	
S1	0	1,167	0,857	S1	0	1,005	0,669	0,996
S2	1,167	0	0,761	S2	1,005	0	1,142	0,583
S3	0,857	0,761	0	S3	0,669	1,142	0	0,896
				S4	0,996	0,583	0,896	0

W analizie skupień istotne jest zrozumienie, co reprezentują elementy należące do poszczególnych skupień, czyli scharakteryzowanie ich właściwości i odkrycie, dlaczego należą do tej samej grupy. Autorzy zdecydowali się na dokonanie takiej charakterystyki, biorąc pod uwagę podział wariantu B na większą liczbę, czyli cztery skupienia – wariantu B. Tabela 7 zawiera szczegółowy opis wszystkich grup ze wskazaniem należących do niej kombinacji topnika i drutu topliwego.

Tabela 7. Odległości pomiędzy centrami skupień w wariacie A i B

Numer skupienia	Elementy skupienia	Opis skupienia
1	1–16, 89–97, 186–188, 193, 228, 230–258, 262–268	<ul style="list-style-type: none"> – grupa topników wysokozasadowych (WZB > 2,0), implikuje to wysoką temperaturę topnienia – dla tej grupy topników możliwe jest uzyskanie dużej czystości stopiwa oraz dużej wartości udarności w niskiej temperaturze – skupienie o wysokiej plastyczności złącza spawanego, najmniejszym wydłużeniu oraz stosunkowo dużej zawartości węgla; znikoma zawartość manganu, molibdenu, chromu i niklu – cena na średnim poziomie
2	106, 118, 172–173, 194–203, 223–227	<ul style="list-style-type: none"> – grupa topników zasadowych charakteryzujących się małą zawartością krzemu, ale dużą molibdenu, manganu, chromu i niklu – ze względu na dużą zawartość pierwiastków stopowych używane głównie do spawania stali wysokostopowych – należą tu topniki o niskiej granicy plastyczności oraz wysokim wydłużeniu – topniki droższe od topników ze skupienia 1 oraz 3, ale tańsze od tych ze skupienia 4
3	17–40, 70–88, 98–100, 166–171, 174–178, 204–220, 229, 259–261, 269–274	<ul style="list-style-type: none"> – grupa topników kwaśnych (mała wartość wskaźnika Boniszewskiego) – duża zawartość węgla i krzemu w spoinie oraz wysoka granica plastyczności i małe wydłużenie – spoina powstała po spawaniu zawiera sporo wtrąceń zmniejszających udarność – najtańsza grupa topników
4	41–46, 48–53, 101–105, 107–117, 119–120, 179–185, 189–192, 221–222	<ul style="list-style-type: none"> – grupa topników obojętnych charakteryzujących się, podobnie jak skupienie 3, niską temperaturą topnienia – ich stosowanie umożliwia uzyskanie bardzo dobrych właściwości spawalniczych, ale pozostawiają też więcej wtrąceń, które zmniejszają możliwość osiągnięcia udarności – mała zawartość węgla w spoinie, a duża krzemu, manganu i chromu – podobnie jak topniki ze skupienia 2, mogą być stosowane do spawania stali wysokostopowych, zauważalne jest osiągnięcie duże wydłużenie – grupa o najwyższej cenie za kg

4. PODSUMOWANIE

W artykule przedstawiono podejście do wspomagania decyzji na temat zakupu materiałów dodatkowych na potrzeby procesu spawania metodą SAW. Analiza została przeprowadzona metodą k -średnich z uwzględnieniem 212 kombinacji topnik–drut topliwy opisanych dziesięcioma kryteriami. W ramach analizy porównano dwa sposoby doboru początkowych centrów skupień, co spowodowało, że w jednym wariancie wygenerowano trzy, a w drugim – cztery skupienia. Następnie szczegółowo opisano grupy wyodrębnione za pomocą wybranego wariantu.

Na podstawie badań wykazano, że analiza skupień może być przydatna we wspomaganiu decyzji zakupowych w branży spawalniczej. Istotne jest jednak prawidłowe przygotowanie danych do analizy, m.in. usunięcie obserwacji odstających, zapewnienie rzetelności danych i odpowiedni wybór kryteriów. Analiza skupień jest na to bardzo czuła.

Autorzy będą kontynuowali i rozszerzali badania biorąc pod uwagę inne metody grupowania. Umożliwi to porównanie uzyskanych skupień, a co za tym idzie – poszerzenie wiedzy o analizowanym problemie.

LITERATURA

1. Abbas, A. (2008). Comparisons between data clustering algorithms. *International Arab Journal of Information Technology*, 5, 320-325.
2. Berkhin, P. (2006). *A Survey of Clustering Data Mining Techniques. Grouping Multidimensional Data*. Springer: Berlin–Heidelberg.
3. Cao, L. (2010). Domain-driven Data Mining: challenges and prospects. *Knowledge and Data Engineering. IEEE Transactions*, 22, 755-769.
4. Houldcroft, P. (2014). *Submerged-arc welding*. Woodhead Publishing.
5. Han, J., Kamber, M. (2006). *Data Mining. Concepts and Techniques*. Morgan Kaufmann.
6. Hastie, T., Tibshirani, R., Friedman, J. (2001). *The elements of statistical learning. Data Mining, inference and prediction*. Springer.
7. Jain, A.K., Murty, M.N., Flynn, P. J. (1999). Data clustering. A review. *ACM Computing surveys (CSUR)*, 31, 264-323.
8. Murtagh, F., Contreras, P. (2012). Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Review: Data Mining and Knowledge Discovery*, 2, 86-97.
9. Popat, S., Emmanuel, M. (2014). Review and comparative study of clustering techniques. *International Journal of Computer Science and Information Technologies*, 5, 805-812.

**PURCHASE DECISION-MAKING SUPPORT IN THE WELDING INDUSTRY
WITH THE USE OF THE K-MEANS METHOD**

Summary

Decision-making in enterprises is often connected with selecting the best solution on the basis of many criteria describing the analyzed problem. From this point of view, it can be called a multi-criterial decision-making problem. The article presents the use of a chosen clustering method – the k-means method – in the selection of materials for the SAW method process. Clusters were divided into two, based on the two different ways of choosing their initial centers. The two options were compared, and finally the clusters created on the basis of the chosen division were characterized in detail. The selected approach proved useful as decision-making support for purchasing materials in the welding industry.

Keywords: decision support, clustering, k-means method